NASA TECHNICAL NOTE

NASA TN D-6642
c.2

NASA TN D-6642

DETECTION OF NONLINEAR TRANSFER FUNCTIONS BY THE USE OF GAUSSIAN STATISTICS

by Jack G. Sheppard

Manned Spacecraft Center
Houston, Texas 77058

1. Report No. TN D-6642	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DETECTION OF NONLINEAR TRANSFER FUNCTIONS BY THE USE OF GAUSSIAN STATISTICS		5. Report Date February 1972
		6. Performing Organization Code
7. Author(s) Jack G. Sheppard, MSC		8. Performing Organization Report No. MSC S-261
		10. Work Unit No. 908-42-07-00-72
9. Performing Organization Name and Address Manned Spacecraft Center Houston, Texas 77058		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Note
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The possibility of using on-line signal statistics to detect electronic-equipment nonlinearities is discussed in this report. Also, the results of an investigation using Gaussian statistics are presented, and a nonlinearity test that uses ratios of the moments of a Gaussian random variable is developed and discussed. Finally, an outline for further investigation is presented.

17. Key Words (Suggested by Author(s)) • Gaussian Statistics • Failure Detection • Linear System(s) • Linear Transfer Function(s)	18. Distribution Statement		
19. Security Classif. (of this report) None	20. Security Classif. (of this page) None	21. No. of Pages 32	22. Price* $3.00

DETECTION OF NONLINEAR TRANSFER FUNCTIONS

BY THE USE OF GAUSSIAN STATISTICS

By Jack G. Sheppard
Manned Spacecraft Center

SUMMARY

Failure detection will be an essential part of electronic systems in future space programs. It would be highly desirable to be able to detect incipient failures while equipment is on-line. The possibility of using the statistics of on-line signals as an indicator of incipient failures is discussed in this report. As a part of this discussion, the concepts of random variables, functions of random variables, and stochastic processes are defined in a limited sense. A nonlinearity test that uses ratios of the moments of a Gaussian random variable is developed and presented. The results of this investigation are encouraging, and the results indicate that further work should be pursued. The next logical step would be to apply nonsupervised learning theory to determine the statistics of nonstationary input signals and to use the results to detect nonlinearities in the output signals of electronic systems.

INTRODUCTION

In the manned spacecraft program, great concern exists about operational reliability; no effort is spared to ensure that, even if the mission cannot be completed satisfactorily, the flight crew is returned to earth safely. One of the foundations of the operational-reliability philosophy is redundancy. Many kinds of redundancy are used (e.g., redundant testing, redundant inspection, redundant functions, and redundant equipment). All these techniques are used to ensure that only reliable equipment is installed in a spacecraft and that, if any item fails, another piece of equipment or mode of operation is available to replace it.

In the past, the greatest emphasis has been placed on exhaustive testing before launch. Usually, the equipment experienced more hours of testing than were experienced in flight. Degraded operation was permitted in the backup modes, but a significant failure caused immediate mission termination. Because of the nature of future manned spacecraft programs, emphasis will be shifted more toward mission-success techniques. For space-station operations, permanence will be emphasized, and space shuttle vehicles with 100-mission lifetimes will spend more time in space than on the ground. Emphasis will be on techniques that minimize ground testing and that avoid mission termination because of equipment malfunction. Degraded backup modes of

$F_X(x), F_Y(y)$	probability that X or Y takes on a value less than or equal to x or y, respectively
$f_X(x), f_Y(y)$	Gaussian density on X or Y
$g_X(x)$	function of X
$h(t)$	function of t
$h'(t)$	first derivative of h with respect to t
i, k	indices or exponents
$\ell(x, t)$	function of x and t
$\ell_t(x, t)$	partial derivative of ℓ with respect to t
$m_i(x)$	i^{th} moment of X
$n = 1, 2, 3, \ldots$	
$P(x_1 \leq x \leq x_2)$	probability that X takes on a value between x_1 and x_2 (or equal to either)
$P(y_1 \leq y \leq y_2)$	probability that Y takes on a value between y_1 and y_2 (or equal to either)
r	dummy variable
t	arbitrary independent variable
V_i	peak value of the ith harmonic in a signal
v_{in}	signal applied to a transfer function
v_{out}	signal received from a transfer function
X	random variable
x_i	possible value of the random variable X
Y	transformation of X
y_i	possible value of Y

3

$\alpha(t)$	function of t
$\alpha'(t)$	first derivative of α with respect to t
$\beta(t)$	function of t
$\beta'(t)$	first derivative of β with respect to t
$\mu_i(x)$	i^{th} central moment of X
σ	standard deviation
σ^2	second central moment or variance

RANDOM VARIABLES

An experiment may be considered in which a marble is selected from a jar containing marbles of different colors. If a number is assigned to each color, then the experiment is as follows.

1. A marble is selected.

2. The color of the marble is determined.

3. The number that corresponds to that color is logged.

Thus, an experiment is performed, an outcome is observed, and a number is assigned to the outcome. This statement is the definition of a random variable (ref. 1). A random variable is a function from a set of outcomes of an experiment to the set of real numbers.

A different experiment might be the selection of a number from the set of all real numbers. In this case, the functional value of the random variable could be the outcome of the experiment itself. This is the type of random variable considered in this report.

FUNCTIONS OF RANDOM VARIABLES

Once a random variable has been defined, the range of the random variable (the set of real numbers associated with the experiment outcomes) can be manipulated by any one of a multitude of functions. The range may be incremented by a constant, multiplied by a constant, subjected to a polynomial transformation, or whatever the imagination can devise. Linear transformations, polynomial transformations, and the densities and moments of random variables will be discussed in this report.

Let x be the value of the random variable X. Then, a linear transformation of X is

$$g_X(x) = a_1 x + a_2 \tag{1}$$

A polynomial transformation of X is

$$g_X(x) = a_1 + a_2 x + a_3 x^2 + \ldots + a_n x^{n-1} \tag{2}$$

The density of a random variable is a function that describes the relative frequency of occurrence of the x-values in the range of the random variable. Many types of densities exist. The uniform density, for example, states that the relative frequency of occurrence of the x-values is constant over a range. This report is concerned with the Gaussian or normal density, which is described by the function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \tag{3}$$

where σ is a spreading factor. The graph of the Gaussian density function is shown in figure 1.

As indicated in figure 1, the x-values tend to cluster around a central point, with a decreasing frequency of occurrence as a function of distance from this point.

The probability that X takes on a value less than or equal to x_1 is called the distribution function and is expressed by
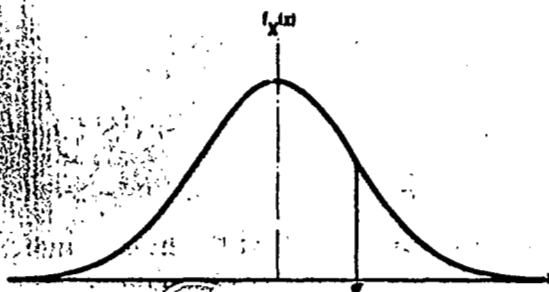

Figure 1. – Gaussian density function.

$$F_X(x) = P\left(x \le x_1\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_1} e^{-x^2/2\sigma^2} dx \tag{4}$$

The density function is the derivative with respect to x of the distribution function (eq. 4).

5

The actual probability that the outcome of the experiment will be between x_1 and x_2 (or equal to either) is found by

$$P\left(x_1 \leq x \leq x_2\right) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-x^2/2\sigma^2} dx \qquad (5)$$

This integral cannot be evaluated by normal means; therefore, numerical techniques must be employed. However, the normal density has been evaluated thoroughly, and tables of values are available in almost any text dealing with statistics (ref. 2).

Moments indicate where the functional values of a random variable are located and how the values are spread in the set of real numbers. The equation for the ith moment of the random variable X is

$$m_i(x) = \int_{-\infty}^{\infty} x^i f_X(x) \, dx \qquad (6)$$

Central moments are formed by subtracting m_1 from each value of x in equation (6); that is,

$$\mu_i(x) = \int_{-\infty}^{\infty} \left(x - m_1\right)^i f_X(x) \, dx \qquad (7)$$

The first moment, called the mean, indicates where the functional values of X are centered in the set of real numbers. The second central moment, called the variance, is usually designated by the symbol σ^2 and indicates how widely dispersed the functional values of X are. Higher order moments also indicate dispersion of the x values.

LINEAR TRANSFORMATION OF GAUSSIAN RANDOM VARIABLES

Let

$$Y = aX \qquad (8)$$

6

where X is a Gaussian random variable. If a is positive, then

$$P\left(y \leq y_1\right) = P\left(x \leq \frac{y_1}{a}\right) \tag{9}$$

or

$$F_Y(y) = F_X\left(\frac{y}{a}\right) \tag{10}$$

If equation (10) is differentiated with respect to y,

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y}{a}\right) \tag{11}$$

If a is negative,

$$P\left(y \leq y_1\right) = P\left(x \geq \frac{y_1}{a}\right) \tag{12}$$

or

$$F_Y(y) = 1 - F_X\left(\frac{y}{a}\right) \tag{13}$$

If equation (13) is differentiated with respect to y,

$$f_Y(y) = -\frac{1}{a} f_X\left(\frac{y}{a}\right) \tag{14}$$

Thus, by combining equations (11) and (14),

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y}{a}\right) \tag{15}$$

or

$$f_Y(y) = \frac{1}{a\sigma\sqrt{2\pi}} e^{-y^2/2a^2\sigma^2} \qquad (16)$$

Two important points become evident as a result of the preceding discussion. First, the linear transformation of a Gaussian random variable results in a Gaussian random variable and, second, the variance of the resulting random variable can be expressed as the variance of the previous random variable multiplied by the square of the transformation constant.

MOMENTS OF A GAUSSIAN RANDOM VARIABLE

Because

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} dx = 1 \qquad (17)$$

therefore,

$$\int_{-\infty}^{\infty} e^{-bx^2} dx = \pi^{1/2} b^{-1/2} \qquad (18)$$

where

$$b = \frac{1}{2\sigma^2} \qquad (19)$$

Leibniz's rule for differentiation states that if

$$h(t) = \int_{\alpha(t)}^{\beta(t)} \ell(x, t) \, dx \qquad (20)$$

8

then

$$h'(t) = \ell[\beta(t), t]\beta'(t) - \ell[\alpha(t), t]\alpha'(t) + \int_{\alpha(t)}^{\beta(t)} \ell_t(x, t)\, dx \qquad (21)$$

By applying Leibniz's rule to equation (18), equation (22) is obtained.

$$\int_{-\infty}^{\infty} -x^2 e^{-bx^2}\, dx = \pi^{1/2}\left(-\frac{1}{2}\right)b^{-3/2} \qquad (22)$$

If equation (18) is differentiated n times with respect to x,

$$\int_{-\infty}^{\infty} x^{2n} e^{-bx^2}\, dx = \pi^{1/2}\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\cdots\left(\frac{2n-1}{2}\right)b^{-(2n+1)/2} \qquad (23)$$

or

$$\int_{-\infty}^{\infty} x^{2n} e^{-bx^2}\, dx = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n}\sqrt{\frac{\pi}{b^{2n+1}}} \qquad (24)$$

with $b = 1/2\sigma^2$

$$\int_{-\infty}^{\infty} x^{2n} e^{-x^2/2\sigma^2}\, dx = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2^n}\sqrt{\frac{\pi}{\left(\frac{1}{2\sigma^2}\right)^{2n+1}}} \qquad (25)$$

or

$$\int_{-\infty}^{\infty} x^{2n} e^{-x^2/2\sigma^2}\, dx = 1 \cdot 3 \cdot 5 \cdots (2n-1)\, \sigma^{2n+1}\sqrt{2\pi} \qquad (26)$$

9

If both sides of equation (26) are divided by $\sigma\sqrt{2\pi}$, then

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x^{2n}e^{-x^2/2\sigma^2}\,dx = 1 \cdot 3 \cdot 5 \cdots (2n-1)\sigma^{2n} \tag{27}$$

which is recognized as the value of the $2n^{th}$ moment of the random variable. Thus, for even moments of a Gaussian random variable,

$$m_i = 1 \cdot 3 \cdot 5 \cdots (2i-1)\sigma^i \tag{28}$$

It can be shown that $m_i = 0$ when i is an odd integer.

High-order moments of a Gaussian random variable can be calculated in the following two ways.

$$m_i = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} x^i e^{-x^2/2\sigma^2}\,dx \tag{29}$$

where i is either an odd or an even integer, or

$$m_i = 1 \cdot 3 \cdot 5 \cdots (2i-1)m_2^{i/2} \tag{30}$$

where i is an even integer. These two methods of calculating the moments suggest a simple way of testing for a nonlinear transfer of Gaussian random variables. In the first method, moments are calculated on the basis of powers of functional values of the random variable. In the second method, moments are calculated on the basis of powers of the second moment. Any consistent shift in density as the functional value moves away from the mean will be emphasized by the higher powers used in calculating higher order moments by the first method. By comparing the moments obtained from the two methods, it should be possible to detect a nonlinear transfer of a Gaussian random variable.

10

STOCHASTIC PROCESSES

A precise definition of stochastic processes has many ramifications. For the purposes of this report, a stochastic process will be considered to be a continuous repetition of the experiment discussed in the section entitled "Random Variables"; that is, at all instants of time, the experiment is being performed, and an outcome is available. Thus, at any instant of time, a random variable exists, and a range value can be obtained as a function of the distribution of that random variable. An example of a stochastic process could be the voltage across the terminals of a battery. If the battery is charged, the voltage has little variation; therefore, the stochastic process is not very interesting. A more interesting stochastic process is the voltage across the terminals of an antenna. This voltage is the sum of various kinds of noise and the different kinds of communications signals currently in use. The contrast between these two kinds of stochastic processes illustrates an important property of some stochastic processes — stationarity. The battery voltage does not change; it remains stationary. The antenna voltage moves about drastically. Actually, the process need not sit still to be stationary. A process is strict-sense stationary if its statistics are not affected by a shift in the time origin. It is wide-sense stationary if its mean is constant and its autocorrelation is a function of the time separation only. Because a Gaussian process is uniquely determined by its first two moments, a wide-sense stationary Gaussian process is also strict-sense stationary and has a constant mean and variance. This report is primarily concerned with Gaussian processes.

GAUSSIAN PROCESS SAMPLING

If the normal signal of some linear system is assumed to be a stationary Gaussian stochastic process, then, at any point in time, the signal value is the functional value of a Gaussian random variable with a fixed mean and variance. Such a random variable from a stochastic process is called a sample. By taking many samples from a process, inferences can be drawn concerning the distribution of the source process. The more samples drawn, the better the inferences will be. In this report, the possibility is investigated that very small amounts of distortion in a supposedly linear transfer of a stationary Gaussian stochastic process can be detected by examination of the statistics of a sample from that process. It is not necessary actually to draw samples from a stochastic process in this investigation. A more convenient method is to use a computer to generate pseudorandom Gaussianly distributed samples, and this was the method used to generate the data presented in this report.

RATIOS OF MOMENTS

A computer program that performs the following operations is described in appendix A.

1. Generation of a predetermined number of Gaussianly distributed random numbers

2. Calculation of m_2 directly

3. Calculation of m_4 directly

4. Evaluation of the ratio $(3m_2^2 - m_4)/3m_2^2$, which ideally should equal zero

5. Repetition of steps 1 to 4 a predetermined number of times

6. Plotting of a histogram of the values obtained from step 4

Results obtained by using this computer program are shown in figure 2. In one case, only 50 numbers were used in the set of random numbers; in another case, 200 numbers were used; and, in a third case, 500 numbers were used. In each case, 1000 values were used in the histogram. As always in sampled statistics, the ideal ratio of zero was not achieved. However, with sample sizes as large as 500, the law of large numbers (ref. 3) applies, and the values are firmly clustered around zero with a relatively small variance. These data formed the basis for a determination of the level of detectable distortion (discussed subsequently), because they represent the ratios of moments to be expected in an undistorted Gaussian process.
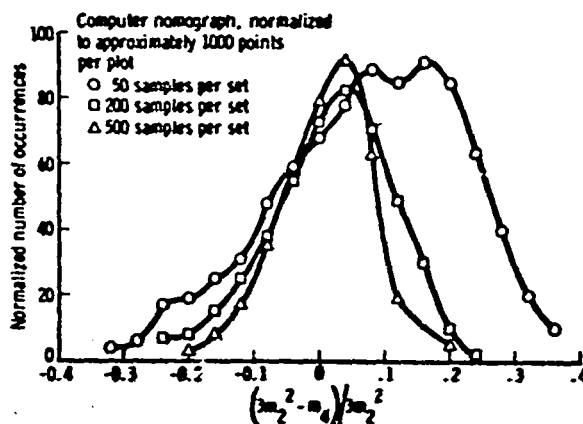

Figure 2. - Sampled standard normal moments for the ratio
$(3m_2^2 - m_4)/3m_2^3$.

In reference 4, information is provided from which the statistics of equation (31) can be calculated.

$$\frac{3m_2^2 - m_4}{3m_2^2} = 1 - \frac{m_4}{3m_2^2} \tag{31}$$

If

$$\tau = \frac{c_4}{c_2^2}\sqrt{\frac{(n-1)(n-2)(n-3)}{24n(n+1)}} \tag{32}$$

where

$$c_2 = \frac{n}{n-1}\, m_2 \tag{33}$$

12

and

$$C_4 = \frac{n^2}{(n-1)(n-2)(n-3)} \left[(n+1)m_4 - 3(n-1)m_2^2 \right] \qquad (34)$$

Then

$$r = \sqrt{\frac{(n-1)(n-2)(n-3)}{24n(n+1)}} \left[\frac{n^2(n-1)^2}{n^2(n-1)(n-2)(n-3)} \right] \left[\frac{(n+1)m_4 - 3(n-1)m_2^2}{m_2^2} \right] \qquad (35)$$

or

$$r = 6 \left(\frac{m_4}{3m_2^2} - \frac{n-1}{n+1} \right) \qquad (36)$$

Therefore,

$$\frac{m_4}{3m_2^2} = \frac{n-1}{n+1} + \frac{r}{6} \qquad (37)$$

and

$$1 - \frac{m_4}{3m_2^2} = 1 - \frac{n-1}{n+1} - \frac{r}{6} \qquad (38)$$

Also in reference 4, the statistics of r (if the assumption is made that the samples are from a Gaussian population) are listed.

$$\mu_1(r) = 0 \qquad (39)$$

$$\mu_2(r) = 1 - \frac{12}{n} + \frac{88}{n^2} - \frac{532}{n^3} \cdots \qquad (40)$$

13

$$\mu_3(r) = 6\sqrt{\frac{6}{n}}\left(1 - \frac{65}{2n} + \frac{4811}{8n^2} - \frac{136\,605}{16n^3} \cdots\right) \tag{41}$$

$$\mu_4(r) = 3 + \frac{468}{n} - \frac{32\,196}{n^2} + \frac{1\,118\,388}{n^3} \cdots \tag{42}$$

Thus, for very large n (n > 100), the mean of $1 - m_4/3m_2^2$ is zero, and its variance approaches 1.7/n.

DISTORTION

Electrical engineers seem to be divided into two groups. One group is concerned with analog signals that have continuous properties, and this group of engineers speaks in terms of linearity, distortion, and harmonics. The other group is concerned with probability theory and speaks in terms of samples, distributions, and statistics. It is difficult to make a correlation between the two groups. Because, in this report, sample statistics are used to investigate linearity, a bridge between the two groups must be used. An attempt was made to correlate harmonic-distortion levels of sine waves with shifts of the statistics of Gaussian random variables.

Distortion is defined as the percentage of content, in a signal, of the harmonics of a sine wave (ref. 5). If the V_i (where $i = 1, 2, \ldots, n$) are the magnitudes of harmonically related sine waves in a signal, with V_1 being the magnitude of the fundamental signal, the percentage of total harmonic distortion D is

$$D = 100\left(\frac{V_2^2 + V_3^2 + \ldots + V_n^2}{V_1^2}\right)^{1/2} \tag{43}$$

Distortion is produced when a sine wave is subjected to a nonlinear transfer and is usually an undesirable effect that indicates unsatisfactory operation. In figure 3, the droop in the v_{in}/v_{out} transfer curve causes the peaks of the input sine wave to be flattened. If the input were a Gaussian stochastic process of approximately the same size as the sine

Figure 3. - Transfer-function curve.

wave, the input would also experience flattening of the larger values. For the ratio $(3m_2^2 - m_4)/3m_2^2$, this flattening would be reflected in a positive shift in the ratio
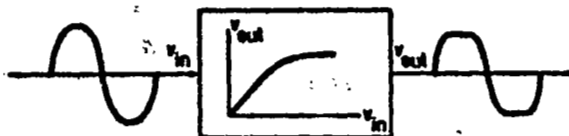
14

obtained. The degree of shift would depend on the amount of distortion. Computer programs written to generate data for correlation of distortion levels with moment-ratio shifts are described in appendixes B and C.

CORRELATION OF SIGNAL MAGNITUDES

Processes "of approximately the same size" were referred to previously; however, this terminology is not very precise. For the purposes of this report, the root mean square (rms) value of a process is used to indicate its "size." For a sine wave, the rms value is 0.707 times the peak value. For nonsinusoidal processes, the rms value must be calculated for each case. The rms value for a zero-mean Gaussian process is the square root of m_2 and is called the standard deviation. Obviously, the form factors of sine waves and Gaussian processes (i.e., the ratios of the peak to the rms values) are different. However, the rms value of any given process represents the same amount of power as the same rms value of any other process. Furthermore, the rms value is linear in a linear transfer; that is, multiplication of a sine wave by a constant has the same effect on the sine-wave rms value as multiplication of a Gaussian process by the same constant has on the standard deviation of the Gaussian process.
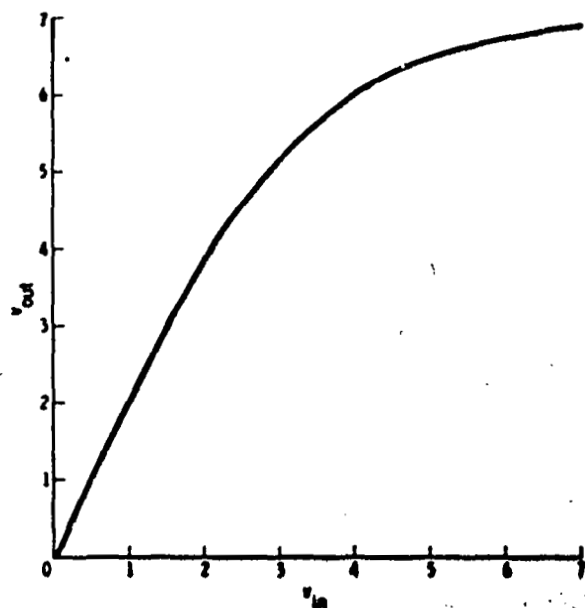
DISTORTION COMPARED WITH TRANSFER FUNCTION

A computer program that calculates the distortion of a sine wave for a given transfer function is described in appendix B. This program performs the following operations.

1. Inputs data that specify the transfer function

2. Fits the least-squares curve to the transfer-function data points

3. Applies a sine wave of a given rms value to the transfer function and obtains the output

4. Constructs a Fourier series on the transferred signal to obtain the harmonics

5. Calculates the distortion

6. Repeats steps 3 to 5 for a different rms value, if desired

Typical transfer functions and distortions obtained for various rms values are shown in figures 4, 5, and 6.
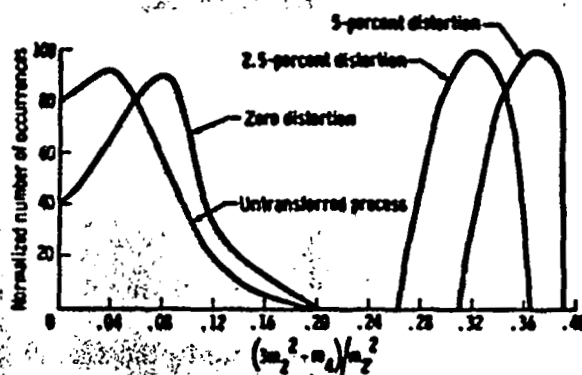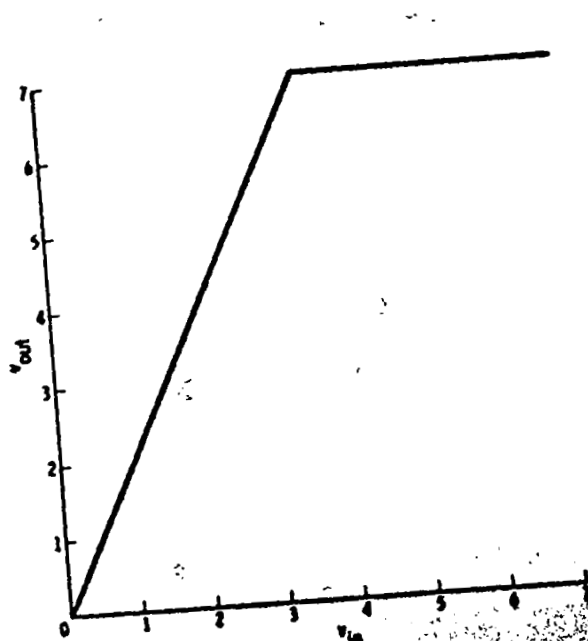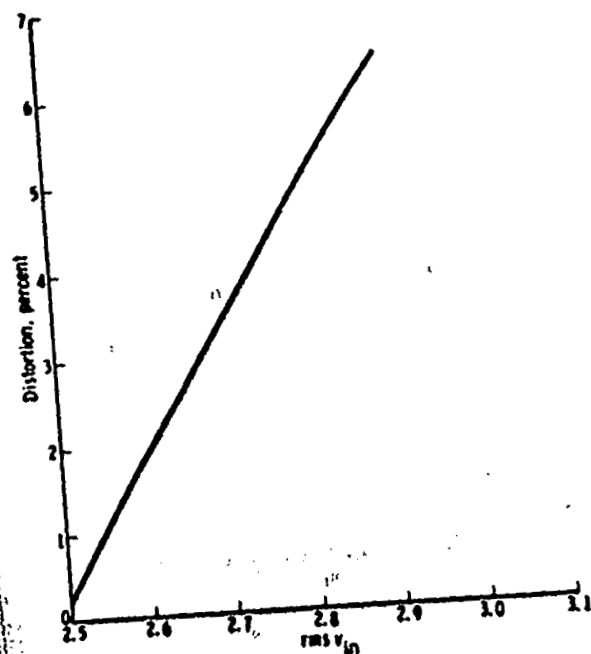
15

(a) Transfer function.
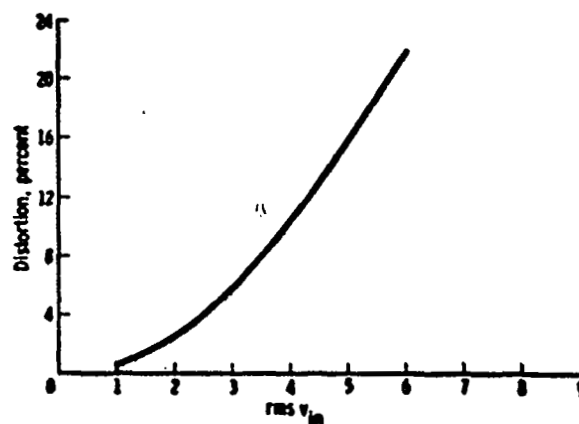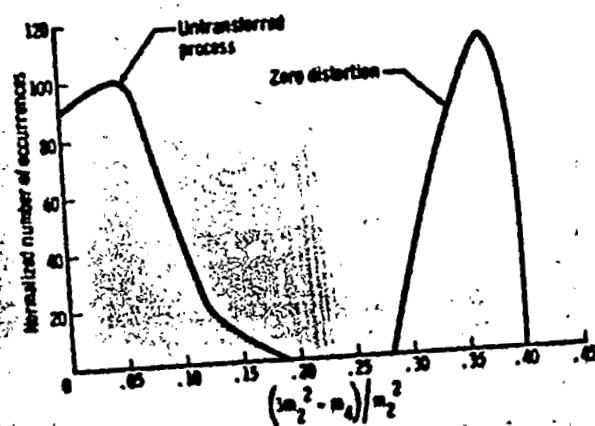

(b) Distortion.

(c) Moment ratios.

Figure 4.- Transfer function 1 ($v_{out} = 1.83347\,v_{in} - 0.0232\,v_{in}^3 + 0.0$, $v_{in} \leq 7$; $v_{out} = 7$, $v_{in} > 7$), distortion for transfer function 1, and moment ratios for transfer function 1.
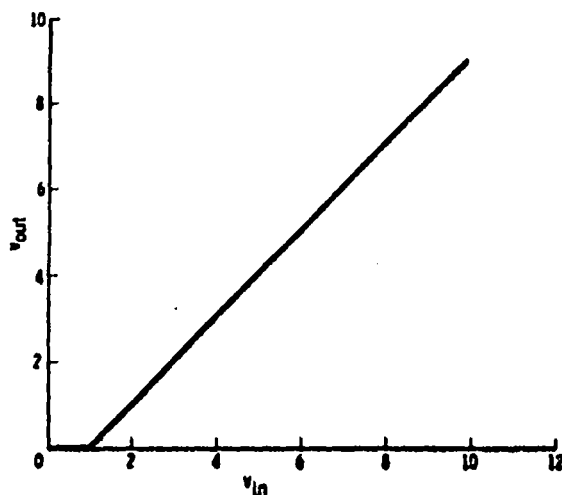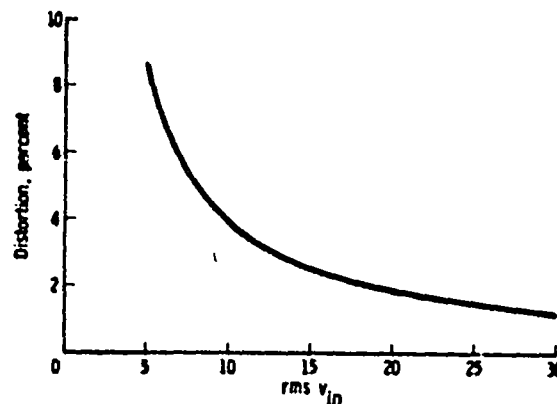
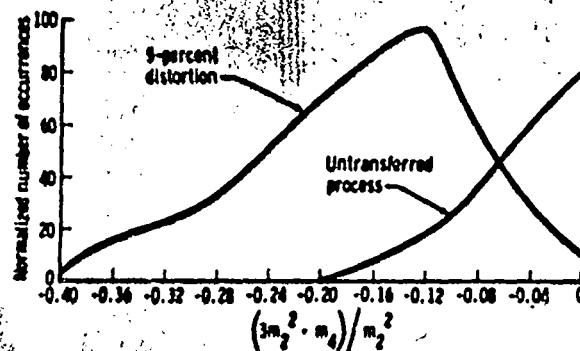(a) Transfer function.

(b) Distortion.

(c) Moment ratios.

Figure 5. - Transfer function 2, distortion for transfer function 2, and moment ratios for transfer function 2.

17

(a) Transfer function.

(b) Distortion.

(c) Moment ratios.

$$\left(3m_2^2 - m_4\right)/m_2^2$$

Figure 6. - Transfer function 3, distortion for transfer function 3, and moment ratios for transfer function 3.

TRANSFERRED GAUSSIAN PROCESSES

A computer program that applies samples from a Gaussian process to transfer functions and examines the moment-ratio shifts is presented in appendix C. This computer program performs the following operations.

1. Inputs the transfer-function polynomial coefficients

2. Generates the random numbers

3. Adjusts the standard deviation of the random numbers to some desired value

4. Applies the random numbers to the transfer function

5. Calculates the first four moments of the transferred process

6. Evaluates the ratio $(3m_2^2 - m_4)/3m_2^2$

7. Repeats as many iterations as desired

8. Plots a histogram of the results of step 6

Histograms produced by the computer program and contrasted with the results of the untransferred Gaussian process are shown in figures 4(c), 5(c), and 6(c).

DISCUSSION OF RESULTS

A typical drooping transfer function, such as might be obtained from a "tired" electronics box, is shown in figure 4(a). This transfer function causes compression of the input-signal peaks and should cause a positive shift in the ratio $(3m_2^2 - m_4)/3m_2^2$. As shown in figure 4(c), such a shift does occur. For relatively small numbers of samples per set, the dispersion of points is so great that small amounts of distortion could not be detected reliably. However, for sample sets as large as 500 (fig. 4(c)), small amounts of distortion can be detected readily. In earlier stages of this investigation, a Kolmogorov-Smirnov goodness-of-fit test was incorporated into the computer program that is described in appendix C. However, this test was eliminated when it became evident that only rarely would a set be rejected as being from a non-Gaussian distribution. The moment-ratio-shift test proved to be much more sensitive, especially because, in the goodness-of-fit test, the sample is assumed to be from a Gaussian process, and the computer must have an extremely good reason before it will reject a sample. In failure detection, the opposite assumption is more desirable, because the penalty for taking a good unit off-line is not high.

The transfer function shown in figure 5(a) is the type that might be obtained in an amplifier with too low a power-supply voltage. The transfer function flattens very sharply and causes a moment-ratio shift in a Gaussian signal that has a rms value that, in a sine wave, would cause no distortion. This difference in distortion is caused by the difference between the form factors of the two signals. The Gaussian signal has no rigid peak value, in theory, and always exceeds the clipping value. For a signal that was usually Gaussian, a standard-deviation value that would not cause excessive clipping of the signal would have to be determined, and that value would be the operating level. Any clipping beyond that value obviously would show up quickly.

The transfer function shown in figure 6(a) is the type obtained when push-push transistors are improperly biased. The transfer function chops out the center of the distribution. Although not as detectable as in the other forms, small amounts of distortion caused by this type of transfer function are still readily detectable.

In summation, the moment-ratio-shift test appears to be an effective way of detecting incipient failures that are reflected in nonlinear transfer functions. The usefulness of this test would be a function of the failure modes and would have to be assessed for each given situation. Nonetheless, the idea of using on-line signal statistics in failure detection appears to have great promise and should be pursued.

SUGGESTIONS FOR FURTHER INVESTIGATION

Although the results presented in the preceding sections are interesting, they point toward the possibility of work of a much broader and more important scope. Most signals used in communications are not Gaussian and are far from stationary. Voice, which is one of the most common types of communications signals, is neither Gaussian nor stationary; voice signals vary in almost every way possible. However, this does not mean that the use of on-line signal statistics as an aid in the detection of incipient failures is unfeasible. Rather, this would appear to be a situation tailormade for the application of learning theory, which is a discipline that has received much attention in recent years.

Nonlinear operations are most likely to occur inside an electronics box. Because the box is probably very small (particularly in a spacecraft), both the input and output signals would be readily available. Furthermore, all future spacecraft will probably have powerful general-purpose computers on board. It seems reasonable to suppose that, in such a situation, the statistics of the input signal to the box could be learned by the computer and compared with the statistics of the output signal of the box. A significant shift in statistics would indicate an incipient failure.

Several aspects of learning theory (ref. 6) will probably be required in such an application. The technique will most certainly require unsupervised learning (ref. 7), because only general characteristics will be known in advance, and the signals will not be stationary. Because of the nonstationary signal characteristics, a form of moving-window technique with optimum stopping rules (ref. 8) will probably be required. Because any given signal has a great variety of statistics, some class of sufficient statistics (ref. 8) must be chosen for manipulation of each signal type.

The most promising method of study in the use of on-line statistics to detect incipient failures appears to be the application of learning theory. Specifically, the following steps should be taken.

1. Investigation of the statistical properties of various communications signals

2. Determination of sufficient statistics, ideally those that are generally applicable in communications signals

3. Application of nonsupervised learning techniques, probably of a moving-window type, in the determination of the sufficient statistics

4. Determination of optimum stopping rules for making "good-bad" decisions about on-line equipment

CONCLUSIONS

The information presented in this report leads to the following conclusions.

1. The nonlinear transfer of a Gaussian signal can be detected by using on-line signal statistics.

2. The moment-ratio-shift test is an effective method for the detection of very small distortion levels.

3. Further investigation is warranted, specifically on the application of learning theory to the problem of detecting incipient failures by using on-line signal statistics.

Manned Spacecraft Center
 National Aeronautics and Space Administration
 Houston, Texas, July 7, 1971
 908-42-07-00-72

REFERENCES

1. Mix, Dwight F.: Random Signal Analysis. Addison-Wesley Pub. Co., 1969.

2. Miller, Irwin; and Freund, John E.: Probability and Statistics for Engineers. Prentice-Hall, Inc., 1965.

3. Papoulis, Athanasios: Probability, Random Variables, and Stochastic Processes. McGraw-Hill Book Co., 1965.

4. Kendall, M. G.; and Stuart, A.: The Advanced Theory of Statistics. Vol. 1. C. Griffin (London), 1958. (Available from Hafner Pub. Co., New York.)

5. Chirlian, Paul M.: Analysis and Design of Electronic Circuits. McGraw-Hill Book Co., 1965.

6. Sheppard, J. G.: Introduction to Learning Machines. NASA TN D-5790, 1970.

7. Fralick, Stanley C.: Learning to Recognize Patterns Without a Teacher. IEEE Transactions on Information Theory, vol. IT-13, no. 1, Jan. 1967, pp. 57-68. (Also available as Stanford Univ. TR 6103-10, SEL-65-011, Mar. 1965.)

8. Fu, K. S.; and Cardillo, G. P.: Optimum Finite Sequential Pattern Recognition. TR-EE67-9, PB-175690, Elec. Eng. Dept., Purdue Univ., 1967.

APPENDIX A

COMPUTER PROGRAM FOR THE DETERMINATION

OF THE RATIOS OF GAUSSIAN MOMENTS

A computer program (fig. A-1) that generates sets of Gaussianly distributed pseudorandom numbers, calculates moments of the sets, and takes ratios of these moments is described in this appendix. A block diagram of the computer program is shown in figure A-2. The operation is described as follows.

1. The control card (first data card) contains the following items.

    a. NNUM — the number of random numbers per set

    b. NRAND — an odd number used to initialize the random-number generator

    c. NRUN — the number of trials to be run

    d. M — the highest order moment to be calculated

2. The instruction $Z = RANDOM (NRAND)$ initializes the canned random-number generator ZOR

3. The DO 6 loop determines the number of trials to be run.

4. The DO 3 loop generates the NNUM random numbers.

5. The canned subroutine CMONTS calculates the first M moments.

6. The remainder of the loop calculates $STOR1 = (3m_2^2 - m_4)/3m_2^2$.

7. The canned subroutine HIS1 calculates and prints a histogram of the NRUN values of STOR1.

Figure A-1. - Computer program for the determination of the ratios of Gaussian moments.

Figure A-2. - Flow diagram of the computer program for the determination of the ratios of Gaussian moments.

APPENDIX B

COMPUTER PROGRAM FOR CALCULATION OF DISTORTION

A computer program (fig. B-1) that calculates the distortion of a sine wave applied to an input transfer function is described in this appendix. The following list is a description of the key data cards that are used in the operation of this computer program.

1. The following control parameters are input on the initial data cards.

    a. N — The number of ordered pairs (X, Y) in a set that defines the transfer function

    b. NP — The number of increments into which 2π is to be divided for the Fourier analysis

    c. NRMS — The number of rms values of the sine waves to be run

    d. KC — The order of the polynomial to be fit to the transfer function

    e. K — The order of the polynomial to be fit to the transfer function

    f. NH — The number of harmonics to be used in the Fourier series

    g. KPOL — The test to see whether data or polynomial coefficients determine the transfer function

        (1) A zero indicates the ordered pair (X, Y) to be read in.

        (2) A one indicates the coefficients to be read in.

    h. CRIT — A critical value (if such exists) above which the transfer function takes on a well-defined value or approaches an asymptote

    i. ASSMPT — A possible asymptotic value of the transfer function

2. Polynomial coefficients (if such exist) are to be read in, seven to a card.

3. Ordered pairs (X, Y) that define a transfer function are to be read in, one pair per card.

4. The rms values to be used are read in, one per card.

A block diagram of this computer program is shown in figure B-2. The computer program performs the following operations.

1. A decision is made on whether data or input polynomial coefficients will define the transfer function, and the choice is read in.

A block diagram of this computer program is shown in figure B-2. The computer program performs the following operations.

1. A decision is made on whether data or input polynomial coefficients will define the transfer function, and the choice is read in.

2. The calling instructions CALL ORTHLS and CALL COEFS call library routines that fit a K-order polynomial to the input data, if such data have been read in.

3. The polynomial coefficients are written.

4. The ordinates of the input data to the transfer function (if such have been read in) and the transferred results are written out.

5. The DO 10 loop sets the values of the abscissas where the Fourier analysis will take place.

6. The DO 20 loop sets the loop that allows iterative operation, with the rms value of a sine wave as the variable.

7. The DO 30 loop (a) sets the peak value of the sine wave at 1.414 times the input rms value and (b) transfers these values through the transfer function (DO 40 loop).

8. The calling instruction CALL DFSRIE calls a library routine that performs a Fourier analysis on the transferred sine wave.

9. The remainder of the computer program calculates the sum of the squares of the nonfundamental terms of the Fourier series and calculates the harmonic distortion.

```
      DIMENSION A(30),Y(30),D(30),C(30),ALPHA(30),BETA(30),TI(30),
     1T2(30),T3(30),A(30),POL(30),B(30),TGRD(30),ABV(30),GRB(30)
      READ(5,1) N,NP,NRMS,CC,E,AN,KPOL,CRIT,ASRPT
    1 FORMAT(7I5,2F5.2)
      A1=E+1

      IF(KPOL) 201,201,208
  208 READ(5,91) (POL(I),I=1,KI)
      GO TO 202
  201 CONTINUE
      DO 2 I=1,N
      READ(5,3) A(I),Y(I)
    3 FORMAT(2F10.5)
    2 CONTINUE
      J=0
      L=0
      CALL ORTHLS(A,Y,N,L,J,C,ALPHA,BETA,E,T1,T2,T3,IND1)
      CALL COEFS(J,C,ALPHA,BETA,AC,E,T1,T2,T3,IND2)
      DO 5 I=1,KI
      POL(I)=C4(I)
    5 CONTINUE
  102 CONTINUE
  101 FORMAT(29H POLYNOMIAL COEFFICIENTS)
      WRITE(6,7) (POL(I),I=1,KI)
      WRITE(6,37)
    7 FORMAT(7F10.5)
      IF(KPOL) 301,301,300
  301 CONTINUE
      WRITE(6,102)
  102 FORMAT(22H INPUT ORDINATES)
      WRITE(6,91) (Y(I),I=1,N)
      WRITE(6,37)
    6 CONTINUE
      DO 70 I=1,N
   14 TGRD(I)=0.
      DO 80 J=1,KI
      TGRD(I)=TGRD(I)+POL(J)+A(I)+(J-1)
   80 CONTINUE
   70 CONTINUE
      WRITE(6,103)
  103 FORMAT(22H TRANSFERRED ORDINATES)
      WRITE(6,91) (TGRD(I),I=1,N)
      WRITE(6,37)
   91 FORMAT(7F10.5)
  300 CONTINUE
      DO 10 I=1,NP
      AP=NP
      A1=I-1
      ABV(I)=3.1416*((I2-E1)/AP)-1.)
   10 CONTINUE
      DO 20 I=1,NRMS
      READ(5,21) RMS
```

```
   21 FORMAT(7F10.5)
      DO 30 J=1,NP
      GRB(J)=0.+1.+SIN(ABV(J))+RMS
      TEST=ABS(GRB(J))
      IF(TEST-CRIT) 500,501,501
  501 Y(J)=ASRPT
      IF(GRB(J).LT. 0.0) Y(J)=-Y(J)
      GO TO 502
  500 CONTINUE
      IF(GRB(J)) 22,23,22
   22 Y(J)=POL(I)

      GO TO 24
   23 Y(J)=0.
      GO TO 502
   24 CONTINUE
      DO 40 K=2,KI
      Y(J)=Y(J)+POL(K)+GRB(J)++(K-1)
   40 CONTINUE
  502 CONTINUE
   30 CONTINUE
      WRITE(6,104)
  104 FORMAT(22H TRANSFERRED SINE WAVE)
      WRITE(6,75) (Y(J),J=1,NP)
      WRITE(6,37)
      ANM=NM+1
      CALL OFSRIE(NP,NM,Y,A,B)
      WRITE(6,105)
  105 FORMAT(23H FOURIER SERIES A TERMS)
      WRITE(6,75) (A(J),J=1,KNM)
      WRITE(6,37)
      WRITE(6,106)
  106 FORMAT(22H FOURIER SERIES B TERMS)
      WRITE(6,75) (B(J),J=1,KNM)
      WRITE(6,37)
      DIST=0.
      DO 60 J=3,KNM
      DIST=DIST+((A(J)++2)+(B(J)++2))
   60 CONTINUE
      PDIST=SQRT(DIST/((A(2)++2)+(B(2)++2)))
   75 FORMAT(7F10.5)
      WRITE(6,107)
  107 FORMAT(19H RMS      DISTORTION)
      WRITE(6,76) RMS,PDIST
   76 FORMAT(2F10.5)
      WRITE(6,38)
   37 FORMAT(1H0)
   38 FORMAT(1H1)
   20 CONTINUE
      STOP
      END
```

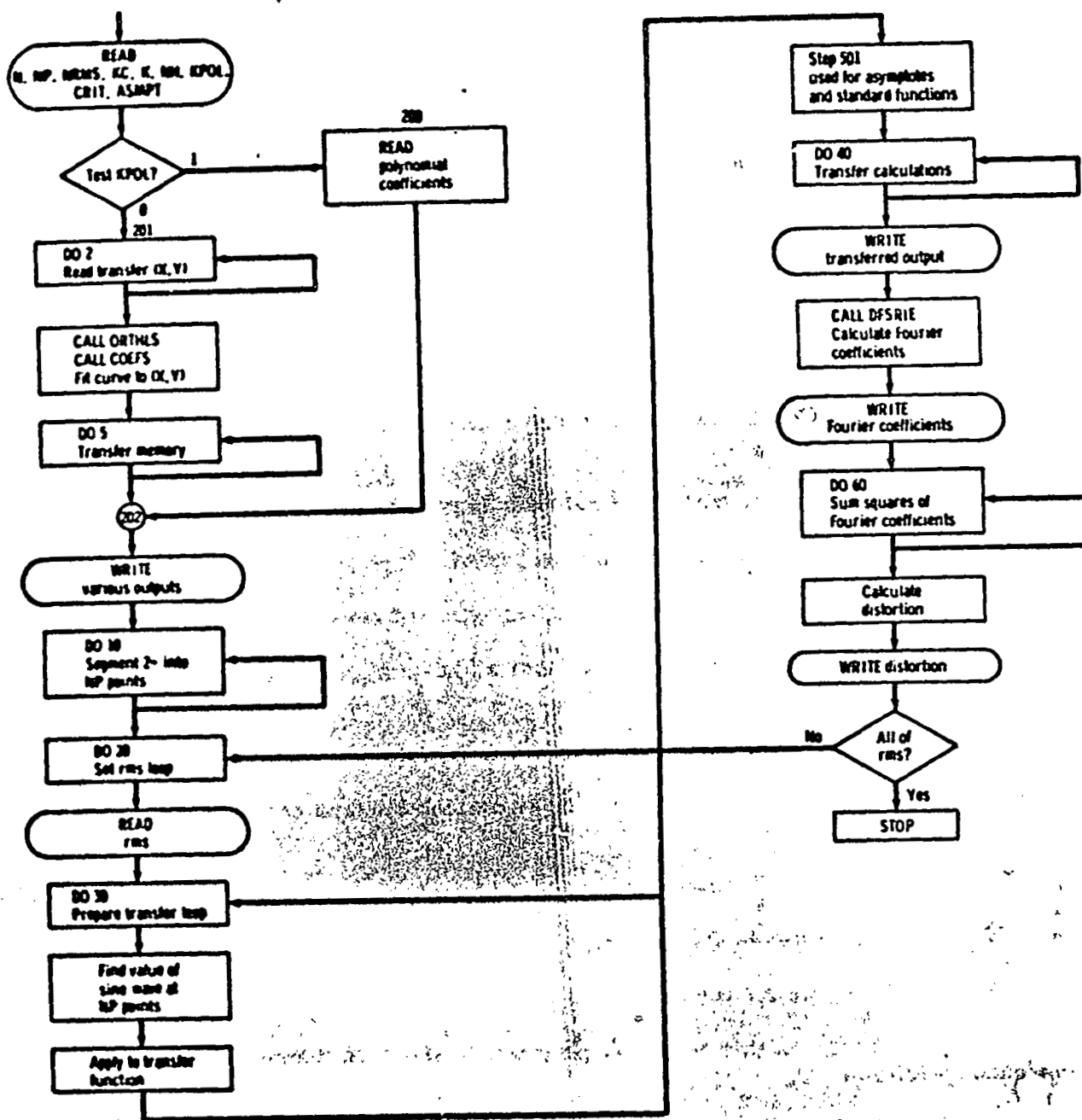Figure B-1. - Computer program for calculation of distortion.

Figure B-2.- Flow diagram of the computer program for calculation of distortion.

APPENDIX C

COMPUTER PROGRAM FOR CALCULATION

OF DISTORTED GAUSSIAN MOMENTS

A computer program (fig. C-1) that generates Gaussianly distributed pseudorandom numbers, applies the numbers to a transfer function, and examines the ratios of moments of the result is described in this appendix. The following is a list of the key data cards used in the operation of this computer program.

1. Initial data card

    a. NPOL — The number of polynomial coefficients to be read in

    b. NNUM — The number of pseudorandom numbers per set

    c. NRAND — An odd number that is used to initialize the random-number generator

    d. NRUN — The number of sets of numbers to be exercised

    e. M — The highest order moment used

    f. NRMS — The number of different standard deviations to be used

    g. INIT — A flexibility number that allows different groups of random numbers to be used from a set

    h. INCR — The increment value for a DO loop

    i. CRIT — The critical value beyond which the transfer function is well defined

    2. POL values — five per card

    3. The rms values — five per card

    4. TITL1 — title for the histogram

A block diagram of this computer program is shown in figure C-2. The program operates as follows.

1. The initial data card is read.

2. The number generator is initialized by $Z = RANDOM (NRAND)$.

3. The polynomial coefficients are read.

4. The rms values are read.

5. The DO 6 loop sets the outer loop for the number of number sets.

6. The DO 3 loop generates the number sets.

7. The DO 7 loop sets the loop to adjust the standard deviation.

8. The DO 8 loop performs the transfer of numbers.

9. The DO 10 loop calculates the moments and ratios of moments.

10. The subroutine HIS1 arranges and prints a histogram of the results.

Figure C-1. - Computer program for calculation of distorted Gaussian moments.

READ
NPOL, NNUM, NRAND, NRUN,
M, NRMS, INIT, INCR, CRIT

Initialize
random-number generator

READ
polynomial coefficients

READ
rms values

DO 6
Outer loop

DO 3
Generate random numbers

DO 7
Set rms values loop

DO 8
Set transfer loop

X < CRIT? No

Yes

Transfer to
standard value

DO 9
Perform transfer

19

All
points
in set? No

Yes

DO 10
Moments loop

CALL CMONTS (X)
Canned program calculates
first M moments

CALL CMONTS (XTRAN)
after transfer moments

DO 31
Calculate $(3m_2^2 - m_4)/3m_2^2$
before and after transfer

WRITE
before and after moment ratios

transferred
STOR1 = moment
ratio

All
rms
values? No

Yes

All
sets? No

Yes

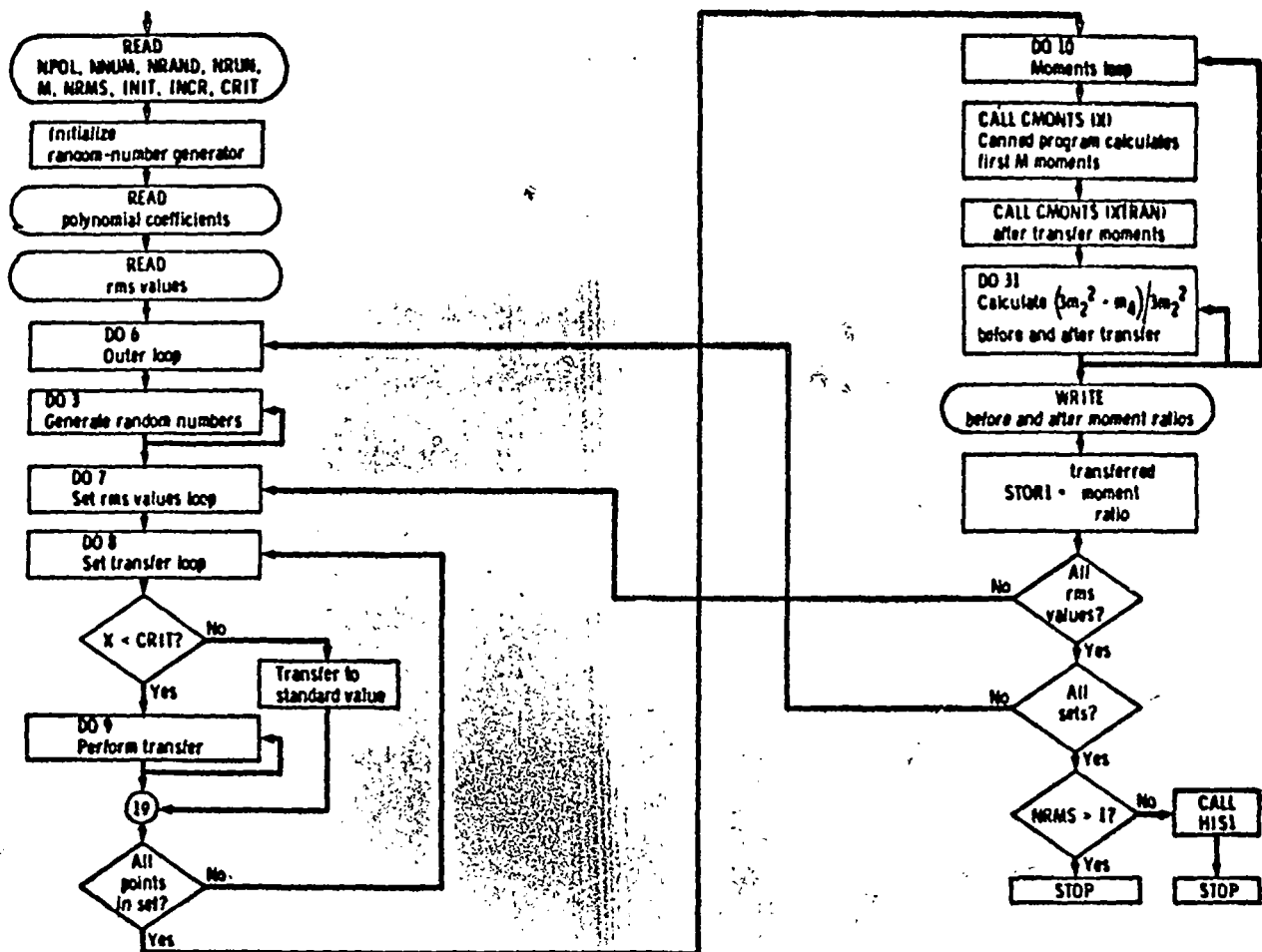NRMS > 1? No   CALL
HIS1

Yes

STOP   STOP

Figure C-2. - Flow diagram of the computer program for calculation of
distorted Gaussian moments.